

A regularity class for the roots of non-negative functions

Kolyan Ray* and Johannes Schmidt-Hieber

Leiden University

Abstract

We investigate the regularity of the positive roots of a non-negative function of one-variable. A modified Hölder space \mathcal{H}^β is introduced such that if $f \in \mathcal{H}^\beta$ then $f^\alpha \in C^{\alpha\beta}$. This provides sufficient conditions to overcome the usual limitation in the square root case ($\alpha = 1/2$) for Hölder functions that $f^{1/2}$ need be no more than C^1 in general. Using a wavelet decomposition of f^α , we derive norm bounds that quantify this relationship.

AMS 2010 Subject Classification: Primary 26A16; secondary 26A27.

Keywords: Regularity of roots; Hölder spaces; differential inequalities; wavelets.

1 Introduction

We study the smoothness properties of roots of non-negative Hölder continuous functions on $[0, 1]$. More precisely, we derive sufficient flatness-type conditions such that if f is in a Hölder space with index β , then f^α has Hölder regularity $\alpha\beta$ for $0 < \alpha \leq 1$.

This question was first studied in the square-root case ($\alpha = 1/2$) by Glaeser [9] and Dieudonné [6], who showed that any non-negative, twice continuously differentiable function on \mathbb{R} admits a continuously differentiable square root. This result is sharp in general in the sense that there exist infinitely differentiable functions such that no admissible square root has Hölder index larger than one (Theorem 2.1 in [1]). Flatness conditions thus become necessary for $\beta > 2$ in order to ensure that a β -smooth function has an admissible square

*The research leading to these results has received funding from the European Research Council under ERC Grant Agreement 320637.

Email: k.m.ray@math.leidenuniv.nl, schmidtthieberaj@math.leidenuniv.nl

root with Hölder index greater than one. Some extensions of the result of Glaeser can be found in Lengyel [12] and Mandai [14]. Macchia [13] showed what regularity $f^{1/r}$ has in a neighbourhood of a zero of f , provided that enough derivatives of f vanish at that point. Some improvements in this direction are found in Reichard [22]. A recent contribution to this question as a subcase of a more general problem can be found in Rainer [20].

Without additional constraints, the (positive) square root of a non-negative and infinitely differentiable function on $[0, 1]$ can be not more than Lipschitz, as the example $f(x) = (x - \frac{1}{2})^2$ shows. However, if a function is flat at all its zeros, in the sense that whenever f becomes small then so do its derivatives, then the square root of f is much more regular. This is the situation we study presently.

There has been recent interest in this question: in a series of papers, Bony *et al.* [1, 2, 3] derive flatness conditions for *admissible* square roots. Recall that g is an admissible square root of f if $f = g^2$ (allowing g to switch signs). It is often possible to find an admissible square root that is much smoother than the positive square root of f . To see this consider again the function $f(x) = (x - \frac{1}{2})^2$. In this case there exists an infinitely differentiable admissible square root, whereas the positive square root is only Lipschitz.

The additional freedom to pick an admissible square root does not always help however. Suppose for instance that f possesses a converging sequence of local minima $(x_n)_n$ such that $f(x_n)$ is strictly positive for all n , but $f(x_n) \rightarrow 0$ as $n \rightarrow \infty$. The non-negative square root is then the most regular admissible square root and there exist infinitely differentiable functions of this form with no admissible square root having Hölder index larger than one [1]. In this article we restrict to studying roots of non-negative functions that do not change sign, which we may without loss of generality take to be positive. We remark that the best situation occurs when f is uniformly bounded away from zero, in which case the Hölder indices of f^α and f coincide (see Lemma 3).

The main result of the article is concisely stated below, with $\|\cdot\|_{C^\beta([0,1])}$ and $\|\cdot\|_{\mathcal{H}^\beta([0,1])}$ denoting the usual and “flat” Hölder norms on $[0, 1]$ respectively, both of which are defined in Section 2.

Theorem 1. *For $\alpha \in (0, 1]$, $\beta > 0$ such that $\alpha\beta \notin \mathbb{N}$, and all non-negative $f \in \mathcal{H}^\beta([0, 1])$,*

$$\|f^\alpha\|_{C^{\alpha\beta}([0,1])} \leq C(\alpha, \beta) \|f\|_{\mathcal{H}^\beta([0,1])}^\alpha.$$

This follows directly from the Theorem 6 below. In particular we see that $f \in \mathcal{H}^\beta$ is a sufficient condition for $\sqrt{f} \in C^{\beta/2}$, thereby overcoming the usual limitation that \sqrt{f} need be no more than C^1 in general. Bounds on the (up to) first-order Hölder or Sobolev norms of $f^{1/r}$ can be found in Glaeser [9] ($r = 2$), Colombini and Lerner [5] and Ghisi and Gobbino

[7]. For bounds for the roots of more general polynomials see the recent work of Parusiński and Rainer [16, 17, 18].

The condition $f \in \mathcal{H}^\beta$ requires a certain notion of “flatness” on the derivatives of f that is different from that considered in [1, 2, 3]. In Section 2 we study the Hölder cone of non-negative functions satisfying such a flatness constraint, which has the properties of a seminorm. We show that the flatness constraints are only necessary for Hölder indices larger than two (Theorem 4). In Section 3 we prove our main result, including some more precise bounds on the norm of f^α . The main idea of the proof is to study f^α using a multiresolution analysis of L^2 , leading to bounds on the wavelet coefficients of f^α (Proposition 1). This is a novel approach allowing one to use the wavelet characterization of Besov spaces to establish the regularity of f^α .

An application of this approach is in nonparametric statistics, where one aims to reconstruct a Hölder function f when observing a noisy version of \sqrt{f} (cf. [15]). In the context of this problem, one must necessarily restrict to positive square roots and it is not enough to find flatness conditions ensuring that \sqrt{f} has some regularity: some control of the Hölder norm is also required [21].

2 Hölder cones and basic properties

Throughout the following let $\lfloor \beta \rfloor$ denote the largest integer strictly smaller than β . For $\beta > 0$ and $A \subset \mathbb{R}$, define

$$|f|_{C^\beta(A)} = \sup_{x \neq y, x, y \in A} \frac{|f^{(\lfloor \beta \rfloor)}(x) - f^{(\lfloor \beta \rfloor)}(y)|}{|x - y|^{\beta - \lfloor \beta \rfloor}}$$

and consider the space of β -Hölder continuous functions on A ,

$$C^\beta(A) = \left\{ f : A \rightarrow \mathbb{R} : f^{(\lfloor \beta \rfloor)} \text{ exists, } \|f\|_{C^\beta(A)} < \infty \right\},$$

where

$$\|f\|_{C^\beta(A)} = \|f\|_\infty + \|f^{(\lfloor \beta \rfloor)}\|_\infty + |f|_{C^\beta(A)}$$

and $\|\cdot\|_\infty$ denotes the $L^\infty(A)$ -norm. This space is often denoted by $C^{[\beta], \beta - [\beta]}(A)$. For convenience, we also write $C^\beta = C^\beta([0, 1])$ for the Hölder space on $[0, 1]$.

For $f \in C^\beta$, define $|f|_{\mathcal{H}^\beta}$ to be the infimum over all non-negative real numbers κ satisfying

$$|f^{(j)}(x)|^\beta \leq \kappa^j |f(x)|^{\beta - j}, \quad \forall x \in [0, 1], \quad \forall 1 \leq j < \beta, \quad (2.1)$$

with $|f|_{\mathcal{H}^\beta} = 0$ for $\beta \leq 1$. This can be written concisely as

$$|f|_{\mathcal{H}^\beta} = \max_{1 \leq j < \beta} \left(\sup_{x \in [0,1]} \frac{|f^{(j)}(x)|^\beta}{|f(x)|^{\beta-j}} \right)^{1/j} = \max_{1 \leq j < \beta} \left\| |f^{(j)}|^\beta / |f|^{\beta-j} \right\|_\infty^{1/j}.$$

The quantity $|f|_{\mathcal{H}^\beta}$ measures the flatness of a function near zero. In particular, if f vanishes at some point, then so do all its derivatives. Define

$$\|f\|_{\mathcal{H}^\beta} = \|f\|_{C^\beta([0,1])} + |f|_{\mathcal{H}^\beta}$$

and set

$$\mathcal{H}^\beta = \{f \in C^\beta : f \geq 0, \|f\|_{\mathcal{H}^\beta} < \infty\}.$$

This space is non-empty as it contains for example the constant functions, any function bounded away from 0 and those of the form $f(x) = (x - x_0)^\beta g(x)$ for $g \geq \varepsilon > 0$ infinitely differentiable.

Recall that if V is a vector space, then $A \subset V$ is called a *convex cone* if $\lambda v \in A$ for all $v \in A$, $\lambda > 0$ and $v, w \in A$ implies $v + w \in A$. In a slight abuse of notation, we say that $\|\cdot\|$ is a norm on A if $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$, $\|\lambda v\| = \lambda \|v\|$ for all $\lambda > 0$, and $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in A$. Thus for a norm on a convex cone, the absolute homogeneity is replaced by positive homogeneity. Similarly, we say that $|\cdot|$ is a seminorm on A if $|v| \geq 0$ for all $v \in A$ and both positive homogeneity and the triangle inequality hold on A .

Theorem 2. *The space \mathcal{H}^β is a convex cone on which $\|\cdot\|_{\mathcal{H}^\beta}$ defines a norm. Moreover, $\|fg\|_{\mathcal{H}^\beta} \leq C(\beta)\|f\|_{\mathcal{H}^\beta}\|g\|_{\mathcal{H}^\beta}$ for any $f, g \in \mathcal{H}^\beta$.*

Proof. For the first part, it is enough to prove that $|f|_{\mathcal{H}^\beta}$ is a seminorm. Observe that $|af|_{\mathcal{H}^\beta} = a|f|_{\mathcal{H}^\beta}$ for $a > 0$. The triangle inequality $|f + g|_{\mathcal{H}^\beta} \leq |f|_{\mathcal{H}^\beta} + |g|_{\mathcal{H}^\beta}$ can be restated as

$$|f^{(j)}(x) + g^{(j)}(x)| \leq (|f|_{\mathcal{H}^\beta} + |g|_{\mathcal{H}^\beta})^{\frac{j}{\beta}} (f(x) + g(x))^{\frac{\beta-j}{\beta}}, \quad \text{for all } x \in [0, 1], \quad j = 1, \dots, [\beta].$$

We may assume that $f(x), g(x), |f|_{\mathcal{H}^\beta}, |g|_{\mathcal{H}^\beta} > 0$, since otherwise the result is trivial. With $r(x) := f(x)/(f(x) + g(x))$ and $\gamma := |f|_{\mathcal{H}^\beta}/(|f|_{\mathcal{H}^\beta} + |g|_{\mathcal{H}^\beta})$, and applying Jensen's inequality to the concave function $x \mapsto x^{\frac{j}{\beta}}$,

$$\begin{aligned} |f^{(j)}(x) + g^{(j)}(x)| &\leq |f|_{\mathcal{H}^\beta}^{\frac{j}{\beta}} f(x)^{\frac{\beta-j}{\beta}} + |g|_{\mathcal{H}^\beta}^{\frac{j}{\beta}} g(x)^{\frac{\beta-j}{\beta}} \\ &= \left[r(x) \left(\frac{\gamma}{r(x)} \right)^{\frac{j}{\beta}} + (1 - r(x)) \left(\frac{1 - \gamma}{1 - r(x)} \right)^{\frac{j}{\beta}} \right] (|f|_{\mathcal{H}^\beta} + |g|_{\mathcal{H}^\beta})^{\frac{j}{\beta}} (f(x) + g(x))^{\frac{\beta-j}{\beta}} \\ &\leq (|f|_{\mathcal{H}^\beta} + |g|_{\mathcal{H}^\beta})^{\frac{j}{\beta}} (f(x) + g(x))^{\frac{\beta-j}{\beta}}. \end{aligned}$$

For the second statement, we establish that each term in the norm $\|fg\|_{\mathcal{H}^\beta}$ is bounded by a constant times $\|f\|_{\mathcal{H}^\beta}\|g\|_{\mathcal{H}^\beta}$. For $\beta \in (0, 1]$ the result follows immediately, so assume that $\beta > 1$. Using that $x^{1/\beta} + y^{1/\beta} \leq 2^{1-1/\beta}(x+y)^{1/\beta}$ for any $x, y \geq 0$,

$$\begin{aligned} |(fg)^{(j)}(x)|^\beta &= \left| \sum_{r=0}^j \binom{j}{r} f^{(r)}(x) g^{(j-r)}(x) \right|^\beta \\ &\leq \left| \sum_{r=0}^j \binom{j}{r} |f|_{\mathcal{H}^\beta}^{\frac{r}{\beta}} |f(x)|^{\frac{\beta-r}{\beta}} |g|_{\mathcal{H}^\beta}^{\frac{j-r}{\beta}} |g(x)|^{\frac{\beta-j+r}{\beta}} \right|^\beta \\ &= \left(|f|_{\mathcal{H}^\beta}^{\frac{1}{\beta}} |g(x)|^{\frac{1}{\beta}} + |g|_{\mathcal{H}^\beta}^{\frac{1}{\beta}} |f(x)|^{\frac{1}{\beta}} \right)^{j\beta} |f(x)g(x)|^{\beta-j} \\ &\leq \left(2^{\beta-1}(|f|_{\mathcal{H}^\beta}\|g\|_\infty + |g|_{\mathcal{H}^\beta}\|f\|_\infty) \right)^j |f(x)g(x)|^{\beta-j}, \end{aligned}$$

from which we deduce that $|fg|_{\mathcal{H}^\beta} \leq 2^{\beta-1}\|f\|_{\mathcal{H}^\beta}\|g\|_{\mathcal{H}^\beta}$ using (2.1) and $\|(fg)^{(\lfloor \beta \rfloor)}\|_\infty \leq 2^{\beta-1}\|f\|_{\mathcal{H}^\beta}\|g\|_{\mathcal{H}^\beta}$. By the triangle inequality and arguing similarly to (3.5) for each term in the sum below, $|(fg)^{(\lfloor \beta \rfloor)}(x) - (fg)^{(\lfloor \beta \rfloor)}(y)|$ is bounded by

$$\begin{aligned} &\sum_{r=0}^{\lfloor \beta \rfloor} \binom{\lfloor \beta \rfloor}{r} \left(|f^{(r)}(x) - f^{(r)}(y)| |g^{(\lfloor \beta \rfloor - r)}(x)| + |g^{(\lfloor \beta \rfloor - r)}(x) - g^{(\lfloor \beta \rfloor - r)}(y)| |f^{(r)}(y)| \right) \\ &\leq C(\beta) \|f\|_{\mathcal{H}^\beta} \|g\|_{\mathcal{H}^\beta} |x - y|^{\beta - \lfloor \beta \rfloor}, \end{aligned}$$

whence $|fg|_{C^\beta} \leq C(\beta) \|f\|_{\mathcal{H}^\beta} \|g\|_{\mathcal{H}^\beta}$. Since $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$, this completes the proof. \square

Remark 1. The space \mathcal{H}^β inherits a notion of completeness from C^β . Suppose that $(f_n)_n \subset \mathcal{H}^\beta$ is a Cauchy sequence in C^β and $\limsup_{n \rightarrow \infty} \|f_n\|_{\mathcal{H}^\beta} < \infty$. Then the C^β -limit f of (f_n) satisfies $f \in \mathcal{H}^\beta$ and $\|f_n\|_{\mathcal{H}^\beta} \rightarrow \|f\|_{\mathcal{H}^\beta}$.

To see the difference between the classical Hölder space C^β and \mathcal{H}^β , consider the functions $f_\gamma : [0, 1] \rightarrow \mathbb{R}$, $f_\gamma(x) = x^\gamma$, $\gamma > 0$. It is well-known that $f_\gamma \in C^\beta$ if and only if either $\gamma \notin \mathbb{N}$ and $\beta \leq \gamma$ or $\gamma \in \mathbb{N}$ and $\beta > \gamma$. The function f_γ thus has arbitrarily large Hölder smoothness if it is a polynomial. One can easily check that for the constrained Hölder spaces, $f_\gamma \in \mathcal{H}^\beta$ if and only if $\beta \leq \gamma$. We next show that the Hölder cones \mathcal{H}^β are nested.

Theorem 3. Let $0 < \beta' \leq \beta$. Then

- (i) $\mathcal{H}^\beta \subset \mathcal{H}^{\beta'}$ and $|f|_{\mathcal{H}^{\beta'}} \leq |f|_{\mathcal{H}^\beta} \vee \|f\|_\infty$,
- (ii) if $f \in \mathcal{H}^\beta$ for $\beta > 1$ and $\inf_x f(x) = 0$, then $|f|_{\mathcal{H}^\beta} \geq \|f\|_\infty$ and $|f|_{\mathcal{H}^{\beta'}} \leq |f|_{\mathcal{H}^\beta}$,
- (iii) if $f \geq c > 0$, then $f \in \mathcal{H}^\beta$ if and only if $f \in C^\beta$.

Proof. (i): It is enough to show that $|f|_{\mathcal{H}^{\beta'}} \leq |f|_{\mathcal{H}^{\beta}} \vee \|f\|_{\infty}$. For $1 \leq j < \beta$, define $I_j^{\beta} := \sup_{x \in [0,1]} (|f^{(j)}(x)|^{\beta} / f(x)^{\beta-j})^{1/j}$ and $x_j^* \in \arg \max_{x \in [0,1]} |f^{(j)}(x)|^{\beta'} / f(x)^{\beta'-j}$. Assume $f(x_j^*) > 0$ since if $f(x_j^*) = 0$, then the result holds trivially as $|f^{(j)}(x_j^*)| = 0$. If $|f^{(j)}(x_j^*)| \leq f(x_j^*)$, then $I_j^{\beta'} \leq \|f\|_{\infty}$. On the contrary, if $|f^{(j)}(x_j^*)| \geq f(x_j^*)$, then $I_j^{\beta'} \leq I_j^{\beta}$. This proves (i).

(ii): By hypothesis, there exists $x_0 \in [0, 1]$ such that $f(x_0) = 0$. By (2.1), this implies that $f^{(j)}(x_0) = 0$ for all $1 \leq j < \beta$, so that $\|f^{(j-1)}\|_{\infty} \leq \|f^{(j)}\|_{\infty}$. This already implies (ii), since for $\beta > 1$,

$$|f|_{\mathcal{H}^{\beta}} \geq \max_{1 \leq j \leq [\beta]} \left(\frac{\|f^{(j)}\|_{\infty}^{\beta}}{\|f\|_{\infty}^{\beta-j}} \right)^{\frac{1}{j}} \geq \|f\|_{\infty}.$$

(iii): Under the assumptions, $|f|_{\mathcal{H}^{\beta}} \leq \max_{1 \leq j < \beta} (c^{j-\beta} \|f^{(j)}\|_{\infty}^{\beta})^{1/j} < \infty$. □

Part (i) of the previous theorem shows that there are two regimes. If the flatness seminorm dominates the L^{∞} -norm, then $|\cdot|_{\mathcal{H}^{\beta}}$ increases in β . This matches our intuition that the seminorms should become stronger for larger β . For functions that are very flat in the sense that $|f|_{\mathcal{H}^{\beta}} \leq \|f\|_{\infty}$, this need not be true. Consider for instance the function $f(x) = x + q$ on $[0, 1]$ with $q > 0$. For $\beta > 1$, $|f|_{\mathcal{H}^{\beta}} = q^{1-\beta}$ and thus for functions of low flatness the seminorm can also decrease in β . The full flatness norms $\|\cdot\|_{\mathcal{H}^{\beta}}$ are however unaffected by this since they also involve $\|f\|_{\infty}$. Statement (ii) of the previous theorem says that the low-flatness phenomenon only occurs if the function is bounded away from zero and in this case (iii) shows that the flatness seminorm is always finite.

For $\beta \in (0, 2]$, the additional derivative constraint is in fact always satisfied and \mathcal{H}^{β} contains all non-negative functions that can be extended to a β -Hölder function on \mathbb{R} .

Theorem 4. *Suppose that $f \in C^{\beta}(\mathbb{R})$ for $\beta \in (0, 2]$ and $f \geq 0$. Let f^* be the restriction of f to $[0, 1]$. Then $f^* \in \mathcal{H}^{\beta}$ and*

$$|f^*|_{\mathcal{H}^{\beta}} \leq 2^{\beta} |f|_{C^{\beta}(\mathbb{R})}.$$

Proof. The result is trivial for $\beta \leq 1$, so assume $\beta \in (1, 2]$. Without loss of generality, we may assume that $|f|_{C^{\beta}(\mathbb{R})} = 1$. Suppose that the statement does not hold, that is $|f^*|_{\mathcal{H}^{\beta}} > 2^{\beta}$. Then there exists $x \in [0, 1]$ such that $|f'(x)| > 2f(x)^{(\beta-1)/\beta}$. By symmetry it is enough to consider $f'(x) > 2f(x)^{(\beta-1)/\beta}$ and we must also have $f(x) > 0$, since otherwise $f'(x) = 0$ as $f \in C^{\beta}(\mathbb{R})$. Set $t = x - f(x)^{1/\beta}$ and observe that for any $y \in [t, x]$,

$$f'(y) \geq f'(x) - |f'(x) - f'(y)| \geq 2f(x)^{(\beta-1)/\beta} - |x - t|^{\beta-1} \geq f(x)^{(\beta-1)/\beta}.$$

Consequently,

$$f(t) = f(x) - \int_t^x f'(s)ds \leq f(x) - f(x)^{(\beta-1)/\beta}|x-t| = 0.$$

Since $f \geq 0$ we must have $f(t) = 0$ but then also $f'(t) = 0 \geq f(x)^{(\beta-1)/\beta} > 0$, which is a contradiction. This proves the claim. \square

Up to smoothness $\beta = 2$, flatness can thus be defined directly without the need to resort to a quantity such as (2.1). This was used in the recent statistical literature, see [19]. To further illustrate the previous result, consider the linear function $f_1(x) = x$. As mentioned above, $f_1 \in C^\beta([0, 1])$ for any $\beta > 0$, but has regularity one with respect to the Hölder cones. Indeed, there is no extension of f_1 to a non-negative function on \mathbb{R} that is smoother than Lipschitz since the non-negativity constraint induces a kink at zero.

The previous theorem cannot be extended to smoothness indices above $\beta = 2$. Indeed, the function $f(x) = (x - 1/2)^2$ is at most in $\mathcal{H}^2([0, 1])$, but can be extended to a function in $C^\beta(\mathbb{R})$ for any $\beta > 0$. The reason is that for smoothness $\beta \in (1, 2]$, any violation of (2.1) must occur at the boundary, since the smoothness and non-negativity constraints together imply (2.1) for interior points. For smoothness indices $\beta > 2$, the given example shows that there can be points in the interior of $[0, 1]$ for which (2.1) does not hold.

A nice feature of classical Hölder spaces is that integration and differentiation respectively increase and decrease the smoothness index by one. Obviously, the Hölder cone \mathcal{H}^β is not closed under differentiation, but the following theorem shows that integration viewed as an operator from \mathcal{H}^β to $\mathcal{H}^{\beta+1}$ is bounded.

Theorem 5. *Given $f \in \mathcal{H}^\beta$, write $F : [0, 1] \rightarrow [0, \infty)$, $F(x) = \int_0^x f(u)du$ for the antiderivative. Then there exists a constant $C(\beta)$ such that*

$$\|F\|_{\mathcal{H}^{\beta+1}} \leq C(\beta)\|f\|_{\mathcal{H}^\beta}.$$

In particular, $F \in \mathcal{H}^{\beta+1}$.

Proof. It is sufficient to prove the result for $\|f\|_{\mathcal{H}^\beta} = 1$. Observe that the Hölder seminorms agree, that is, $|F|_{C^{\beta+1}} = |f|_{C^\beta}$. It remains to show that for some κ which depends only on β , $|F^{(j)}(x)| \leq \kappa^{\frac{j}{\beta+1}} F(x)^{\frac{\beta+1-j}{\beta+1}}$, for all $x \in [0, 1]$ and all $j = 0, \dots, \lfloor \beta \rfloor + 1$. Notice that by Lemma 2, $f(x+h) \geq f(x)/2$ for all $|h| \leq af(x)^{1/\beta}$, for some positive constant $a = a(\beta)$. First, we prove the case $j = 1$. We show that

$$f(x)^{\frac{\beta+1}{\beta}} \leq \frac{2}{a} F(x). \tag{2.2}$$

If $af(x)^{1/\beta} \leq x$, then $f(x)^{\frac{\beta+1}{\beta}} \leq \frac{x}{a}f(x) \leq \frac{2}{a} \int_0^x f(u)du = \frac{2}{a}F(x)$. On the other hand if $af(x)^{1/\beta} \geq x$, then $f(x)^{\frac{\beta+1}{\beta}} \leq \frac{2}{a} \int_{x-af(x)^{1/\beta}}^x f(u)du \leq \frac{2}{a}F(x)$. This proves (2.2).

Recall that $\|f\|_{\mathcal{H}^\beta} = 1$. Together with (2.2), for $j = 2, \dots, \lfloor \beta \rfloor + 1$, $|F^{(j)}(x)| = |f^{(j-1)}(x)| \leq |f(x)|^{\frac{\beta+1-j}{\beta}} \leq (2/a)^{\frac{\beta}{\beta+1}} F(x)^{\frac{\beta+1-j}{\beta+1}}$. This proves the differential inequalities for $j = 2, \dots, \lfloor \beta \rfloor + 1$. With (2.2), the statement follows. \square

3 Norm bounds for f^α and proof of Theorem 1

Let ϕ, ψ be bounded, compactly supported S -regular scaling function and wavelet, respectively, where $S \in \mathbb{N}$; in particular we assume that $\int x^i \psi(x) dx = 0$ for $i = 0, \dots, S-1$ (see e.g. [10] for more details). Let $\{\phi_k : k \in \mathbb{Z}\} \cup \{\psi_{j,k} : j = 0, 1, \dots, k \in \mathbb{Z}\}$ be the corresponding compactly supported wavelet basis of $L^2(\mathbb{R})$. From this we can construct an S -regular wavelet basis of $L^2([0, 1])$ by selecting all basis functions with support intersecting the interval $[0, 1]$ and then correcting for boundary effects as explained in Theorem 4.4 of [4]. Thus, there are sets of indices I_j with cardinality bounded by a multiple of 2^j such that $\{\phi_k : k \in I_{-1}\} \cup \{\psi_{j,k} : j = 0, 1, \dots, k \in I_j\}$ forms an orthonormal basis of $L^2([0, 1])$.

The following theorem states that the function $f \mapsto f^\alpha$ maps \mathcal{H}^β into the Besov space $B_{\infty\infty}^{\alpha\beta}$ (see Chapter 4 of [8] for a full definition). We recall that $B_{\infty\infty}^\beta = C^\beta$ for non-integer β , while $B_{\infty\infty}^\beta$ equals the slightly larger Hölder-Zygmund space for integer β . Contrast this with the square root case $\alpha = 1/2$, which maps the full Hölder space C^β into C^1 , but not $C^{\beta/2}$ for any $\beta > 2$ (Theorem 2.1 of [1]). In view of the continuous embedding $\mathcal{H}^\beta \subset C^\beta$, we note that the following result implies Theorem 1.

Theorem 6. *For $\alpha \in (0, 1]$, $\beta > 0$ and all $f \in \mathcal{H}^\beta$,*

$$\|f^\alpha\|_{B_{\infty\infty}^{\alpha\beta}} \leq C(\alpha, \beta) \|f\|_{\mathcal{H}^\beta}^\alpha,$$

$$|f^\alpha|_{\mathcal{H}^{\alpha\beta}} \leq C(\alpha, \beta) |f|_{\mathcal{H}^\beta}^\alpha.$$

In particular for $\alpha\beta \notin \mathbb{N}$,

$$\|f^\alpha\|_{\mathcal{H}^{\alpha\beta}} \leq C(\alpha, \beta) \|f\|_{\mathcal{H}^\beta}^\alpha.$$

Theorem 6 states that if $f \in \mathcal{H}^\beta$, then f^α has smoothness index $\alpha\beta$. Let us compare this to the flatness conditions considered in the literature on admissible square roots. For a given function f , assume that there exists a continuous function γ , vanishing on the set of flat points of f , such that for any positive minimum x_0 of f , $f''(x_0) \leq \gamma(x_0)f(x_0)^{1/2}$. Theorem

3.5 of Bony et al. [1] establishes that this is a necessary and sufficient condition for a four times continuously differentiable function to have a twice continuously differentiable, admissible square root. This is comparable to the flatness seminorm (2.1), which for $\beta = 4$ also gives that f'' should be bounded by $f^{1/2}$. Nevertheless, in order to obtain bounds on the norm of \sqrt{f} , constraints on the third derivative of f must also be imposed in our framework.

One can exploit extra regularity by assuming that f and its derivatives vanish at all local minima (e.g. in [1, 2, 3] for admissible square roots). However, we take a different approach permitting functions to take small non-zero values. In fact, as mentioned in [1], the obstacle preventing the existence of a twice continuously differentiable square root for a general non-negative four times continuously differentiable function is a converging sequence of relatively small non-zero minima. A sufficient condition for arbitrary even integer $\beta \geq 4$ was found in Theorem 2.2 in [3]: a β -times continuously differentiable function f has a $\beta/2$ -times continuously differentiable square root if at each local minimum of f , the function and its derivatives up to order $\beta - 4$ vanish. While this constraint only needs to hold for the local minima of f , in certain instances it can be more restrictive than the flatness condition (2.1); in particular, it forces all local minima to be exactly zero. For $\beta \geq 4$, consider the following functions on $[0, 1]$:

$$f_\delta(x) = x^\beta + \delta^{\beta-2}x^2 + \delta^\beta, \quad 0 \leq \delta \leq \delta_0,$$

for some fixed $\delta_0 > 0$. Observe that $|f'_\delta(x)| \leq 2\beta(x + \delta)^{\beta-1}$, $|f''_\delta(x)| \leq 2\beta(\beta - 1)(x + \delta)^{\beta-2}$ and $(x + \delta)^\beta \leq 2^\beta f_\delta(x)$. Thus, $\|f_\delta\|_{\mathcal{H}^\beta} \leq C(\beta)(1 + \delta^{\beta-2} + \delta) \leq C(\beta, \delta_0)$, so that applying Theorem 6 yields bounds for $\|\sqrt{f_\delta}\|_{\mathcal{H}^{\beta/2}}$, which are uniform over $\{f_\delta : 0 \leq \delta \leq \delta_0\}$. On the other hand, the condition in Theorem 2.2 in [3] is not satisfied since f_δ and f''_δ do not vanish at the local minimum $x = 0$ for any $\delta > 0$.

The proof of Theorem 6 relies on decomposing f^α using a suitable wavelet basis. We obtain two bounds on the wavelet coefficients $|\langle f^\alpha, \psi_{j,k} \rangle|$, whose decay characterizes the Besov norms $\|f^\alpha\|_{B_{pq}^\beta}$. The first holds for all (j, k) and is most useful for those $\psi_{j,k}$ on whose support the function f is small. The second bound depends explicitly on the local function values and becomes useful when f is large, in which case f^α typically has more regularity than $\alpha\beta$ by Lemma 3.

Proposition 1. *Suppose that $\alpha \in (0, 1]$, ψ is S -regular and that $f \in \mathcal{H}^\beta$ for $0 < \beta < S$. Then*

$$|\langle f^\alpha, \psi_{j,k} \rangle| \leq C(\psi, \alpha, \beta) \|f\|_{\mathcal{H}^\beta}^\alpha 2^{-j(\alpha\beta+1/2)}.$$

For $x_0 \in [0, 1]$, let $j(x_0)$ be the smallest integer satisfying $2^{j(x_0)} \geq |\text{supp}(\psi)|a^{-1}(\|f\|_{\mathcal{H}^\beta}/f(x_0))^{1/\beta}$ where $a = a(\beta)$ is the constant in Lemma 2. Then for any wavelet $\psi_{j,k}$ with $j \geq j(x_0)$ and

$x_0 \in \text{supp}(\psi_{j,k})$,

$$|\langle f^\alpha, \psi_{j,k} \rangle| \leq C(\psi, \alpha, \beta) \frac{\|f\|_{\mathcal{H}^\beta}}{f(x_0)^{1-\alpha}} 2^{-\frac{j}{2}(2\beta+1)}.$$

To prove Proposition 1, and hence Theorem 6, we require three technical lemmas. Lemma 1 can be easily proved by Taylor expanding f using the moment properties of the wavelet function.

Lemma 1. *Suppose that the wavelet function is S -regular. If $f \in C^\beta([0, 1])$ for $0 < \beta < S$, then there exists a function g with $\|g\|_\infty \leq 1$, such that for any $x_0 \in (0, 1)$,*

$$\left| \int f(x) \psi_{j,k}(x) dx \right| \leq \frac{1}{[\beta]!} \left| \int [f^{(\lfloor \beta \rfloor)}(x_0 + g(y)(y - x_0)) - f^{(\lfloor \beta \rfloor)}(x_0)] (y - x_0)^{[\beta]} \psi_{j,k}(y) dy \right|.$$

The wavelet bounds we derive depend crucially on the local function size. The following lemma quantifies this locality by establishing a neighbourhood of each point on which the function is relatively constant.

Lemma 2. *Suppose that $f \in \mathcal{H}^\beta$ with $\beta > 0$ and let $a = a(\beta) > 0$ be any constant satisfying $(e^a - 1) + a^\beta / ([\beta]!) \leq 1/2$. Then for*

$$|h| \leq a \left(\frac{|f(x)|}{\|f\|_{\mathcal{H}^\beta}} \right)^{1/\beta},$$

we have

$$|f(x+h) - f(x)| \leq \frac{1}{2} |f(x)|,$$

implying in particular, $|f(x)|/2 \leq |f(x+h)| \leq 3|f(x)|/2$.

Proof. Without loss of generality, we may assume that $\|f\|_{\mathcal{H}^\beta} = 1$. By a Taylor expansion and the definition of \mathcal{H}^β , there exists ξ between x and $x+h$ such that

$$\begin{aligned} |f(x+h) - f(x)| &= \left| \sum_{j=1}^{[\beta]-1} \frac{f^{(j)}(x) h^j}{j!} + \frac{1}{[\beta]!} f^{(\lfloor \beta \rfloor)}(\xi) h^{[\beta]} \right| \\ &\leq \sum_{j=1}^{[\beta]} \frac{|f(x)|^{\frac{\beta-j}{\beta}} h^j}{j!} + \frac{1}{[\beta]!} \left| f^{(\lfloor \beta \rfloor)}(\xi) - f^{(\lfloor \beta \rfloor)}(x) \right| h^{[\beta]} \\ &\leq \sum_{j=1}^{[\beta]} \frac{a^j}{j!} |f(x)| + \frac{a^\beta}{[\beta]!} |f(x)| \leq \frac{1}{2} |f(x)|. \end{aligned}$$

□

For $f \in \mathcal{H}^\beta$, the function f^α satisfies a Hölder-type condition with exponent β and locally varying Hölder constant. The following is the key technical result for establishing the smoothness of f^α and hence the decay of $|\langle f^\alpha, \psi_{j,k} \rangle|$. The main ingredient in the proof is a careful analysis of Faà di Bruno's formula, which generalizes the chain rule to higher derivatives [11]:

$$\frac{d^k}{dx^k} h(f(x)) = \sum_{(m_1, \dots, m_k) \in \mathcal{M}_k} \frac{k!}{m_1! \dots m_k!} h^{(m_1 + \dots + m_k)}(f(x)) \prod_{j=1}^k \left(\frac{f^{(j)}(x)}{j!} \right)^{m_j}, \quad (3.1)$$

where \mathcal{M}_k is the set of all k -tuples of non-negative integers satisfying $\sum_{j=1}^k j m_j = k$. Note that for $h(x) = x^\alpha$, we have $h^{(r)}(x) = C_{\alpha,r} x^{\alpha-r}$ for some $C_{\alpha,r} \neq 0$ (except the trivial case $\alpha = 1$). We can relate the derivatives appearing in (3.1) to f using the seminorm $|\cdot|_{\mathcal{H}^\beta}$.

Lemma 3. *For $\alpha \in (0, 1]$, $\beta > 0$, there exists a constant $C(\alpha, \beta)$ such that for all $f \in \mathcal{H}^\beta$, $0 \leq k < \beta$ and $x, y \in [0, 1]$,*

$$|(f(x)^\alpha)^{(\lfloor \beta \rfloor)} - (f(y)^\alpha)^{(\lfloor \beta \rfloor)}| \leq \frac{C(\alpha, \beta)(|f|_{C^\beta} + |f|_{\mathcal{H}^\beta})}{\min(f(x)^{1-\alpha}, f(y)^{1-\alpha})} |x - y|^{\beta - \lfloor \beta \rfloor},$$

and

$$\left| \frac{d^k}{dx^k} (f(x)^\alpha) \right| \leq C(\alpha, \beta) \|f\|_{\mathcal{H}^\beta}^{k/\beta} f(x)^{\alpha - k/\beta}. \quad (3.2)$$

Moreover, if $f \geq \varepsilon > 0$,

$$\|f^\alpha\|_{\mathcal{H}^\beta} \leq \frac{C(\alpha, \beta)}{\varepsilon^{1-\alpha}} \|f\|_{\mathcal{H}^\beta}.$$

Proof. Without loss of generality assume $f(y) \leq f(x)$ and $|f|_{C^\beta} + |f|_{\mathcal{H}^\beta} = 1$. For $\beta \in (0, 1]$, by the mean value theorem,

$$|f(x)^\alpha - f(y)^\alpha| \leq \sup_{f(y) \leq t \leq f(x)} \alpha t^{\alpha-1} |f(x) - f(y)| \leq \frac{|x - y|^\beta}{f(y)^{1-\alpha}}.$$

Consider now $\beta > 1$ and write $k = \lfloor \beta \rfloor$ (the following also holds for all $k \leq \lfloor \beta \rfloor$ with certain simplifications). We must consider separately the two cases where $|x - y|$ is small and large. Let $C(\alpha, \beta)$ be a generic constant, which may change from line to line.

Suppose first that $|x - y| \leq a f(x)^{1/\beta}$ with a as in Lemma 2. By Lemma 2 we have $f(y)/2 \leq f(x) \leq 3f(y)/2$, which will be used freely without mention in the following. The proof is based on a careful analysis of Faà di Bruno's formula (3.1).

We shall establish the result by proving a Hölder bound for each of the summands in (3.1) individually. Fix a k -tuple $(m_1, \dots, m_k) \in \mathcal{M}_k$ and write $M := \sum_{j=1}^k m_j$. By the triangle

inequality

$$\begin{aligned}
& \left| f(x)^{\alpha-M} \prod_{j=1}^k \left(f^{(j)}(x) \right)^{m_j} - f(y)^{\alpha-M} \prod_{j=1}^k \left(f^{(j)}(y) \right)^{m_j} \right| \\
& \leq \left| \left(f(x)^{\alpha-M} - f(y)^{\alpha-M} \right) \prod_{j=1}^k \left(f^{(j)}(x) \right)^{m_j} \right| \\
& \quad + \left| f(y)^{\alpha-M} \left(\prod_{j=1}^k \left(f^{(j)}(x) \right)^{m_j} - \prod_{j=1}^k \left(f^{(j)}(y) \right)^{m_j} \right) \right|.
\end{aligned} \tag{3.3}$$

Before bounding the terms in (3.3), we require some additional estimates. Firstly, by the definition of \mathcal{H}^β ,

$$\left| \prod_{j=1}^k \left(f^{(j)}(x) \right)^{m_j} \right| \leq \prod_{j=1}^k |f(x)|^{\frac{(\beta-j)m_j}{\beta}} = |f(x)|^{M-k/\beta}. \tag{3.4}$$

Secondly, for any function g and integer $r \geq 1$, we have by the mean value theorem $g(x)^r - g(y)^r = rg(\xi)^{r-1}g'(\xi)(x-y)$ for some ξ between x and y . Noting that $f(\xi) \approx f(x) \approx f(y)$, that $\beta - k \in (0, 1]$ and applying the above to $g(x) = f^{(j^*)}(x)$ with $r = m_{j^*}$ and $j^* \in \{1, \dots, k-1\}$ yields

$$\begin{aligned}
& \left| \left(f^{(j^*)}(x) \right)^{m_{j^*}} - \left(f^{(j^*)}(y) \right)^{m_{j^*}} \right| \\
& \leq \left| m_{j^*} \left(f^{(j^*)}(\xi_{x,y,j^*}) \right)^{m_{j^*}-1} f^{(j^*+1)}(\xi_{x,y,j^*})(x-y) \right|^{\beta-k} \left| \left(f^{(j^*)}(x) \right)^{m_{j^*}} + \left(f^{(j^*)}(y) \right)^{m_{j^*}} \right|^{1-(\beta-k)} \\
& \leq C(\beta) m_{j^*}^{\beta-k} |f(\xi_{x,y,j^*})|^{\frac{(m_{j^*}(\beta-j^*)-1)(\beta-k)}{\beta}} |x-y|^{\beta-k} |f(y)|^{\frac{m_{j^*}(\beta-j^*)(1-\beta+k)}{\beta}} \\
& \leq C(\beta) |f(y)|^{\frac{m_{j^*}(\beta-j^*)-(\beta-k)}{\beta}} |x-y|^{\beta-k}.
\end{aligned} \tag{3.5}$$

Strictly speaking, we cannot invoke (3.5) for $j^* = k$, since f is only k -times differentiable. However, noting that m_k can only take values 0 or 1, we see directly from the Hölder continuity of $f^{(k)}$ that the conclusion of (3.5) holds as well for $j^* = k$ since we have the bound $|x-y|^{\beta-k}$. By the same argument as (3.5)

$$\begin{aligned}
|f(x)^{\alpha-M} - f(y)^{\alpha-M}| & \leq C(\alpha, \beta) \frac{|f(y)|^{M-1-\alpha+k/\beta} |x-y|^{\beta-k}}{f(x)^{M-\alpha} f(y)^{M-\alpha}} \\
& \leq C(\alpha, \beta) |f(y)|^{-M-1+\alpha+k/\beta} |x-y|^{\beta-k}.
\end{aligned} \tag{3.6}$$

Using (3.4) and (3.6), the first term in (3.3) is bounded by $C(\beta) |f(y)|^{\alpha-1} |x-y|^{\beta-k}$ as required.

For the second term in (3.3), we repeatedly apply the triangle inequality, each time changing the variable in a single derivative. Fix j^* and define vectors $(z_j^{j^*})_{j=1}^k, (\tilde{z}_j^{j^*})_{j=1}^k$ that are identically equal to x or y in all entries and differ only in the j^* -coordinate, where $z_{j^*}^{j^*} = x, \tilde{z}_{j^*}^{j^*} = y$. Using (3.5)

$$\begin{aligned}
& f(y)^{\alpha-M} \left| \prod_{j=1}^k \left(f^{(j)}(z_j^{j^*}) \right)^{m_j} - \prod_{j=1}^k \left(f^{(j)}(\tilde{z}_j^{j^*}) \right)^{m_j} \right| \\
&= f(y)^{\alpha-M} \prod_{j=1, j \neq j^*}^k \left(f^{(j)}(z_j^{j^*}) \right)^{m_j} \left| \left(f^{(j^*)}(x) \right)^{m_{j^*}} - \left(f^{(j^*)}(y) \right)^{m_{j^*}} \right| \\
&\leq C(\alpha, \beta) f(y)^{\alpha-M} \left(\prod_{j=1, j \neq j^*}^k |f(z_j^{j^*})|^{\frac{(\beta-j)m_j}{\beta}} \right) |f(y)|^{\frac{m_{j^*}(\beta-j^*)-(\beta-k)}{\beta}} |x-y|^{\beta-k} \\
&\leq C(\alpha, \beta) |f(y)|^{\alpha-M+\sum_j (\beta-j)m_j/\beta-(\beta-k)/\beta} |x-y|^{\beta-k} \\
&= C(\alpha, \beta) |f(y)|^{\alpha-1} |x-y|^{\beta-k},
\end{aligned} \tag{3.7}$$

where in the last line we have used that $\sum_j j m_j = k$. By repeatedly applying the triangle inequality and using (3.7), we can bound the second term in (3.3) by

$$f(y)^{\alpha-M} \sum_{j^*=1}^k \left| \prod_{j=1}^k \left(f^{(j)}(z_j^{j^*}) \right)^{m_j} - \prod_{j=1}^k \left(f^{(j)}(\tilde{z}_j^{j^*}) \right)^{m_j} \right| \leq C(\alpha, \beta) |f(y)|^{\alpha-1} |x-y|^{\beta-k},$$

thereby completing the proof in the case $|x-y| \leq a f(x)^{1/\beta}$.

Applying Faà di Bruno's formula (3.1) and (3.4) yields

$$\left| \frac{d^k}{dx^k} (f(x)^\alpha) \right| \leq \sum_{(m_1, \dots, m_k) \in \mathcal{M}_k} C f(x)^{\alpha-M} f(x)^{M-k/\beta} \leq C f(x)^{\alpha-k/\beta}. \tag{3.8}$$

For $|x-y| > a f(x)^{1/\beta}$ we thus have

$$\begin{aligned}
|(f(x)^\alpha)^{(k)} - (f(y)^\alpha)^{(k)}| &\leq C(\alpha, \beta) (f(x)^{\alpha-k/\beta} + f(y)^{\alpha-k/\beta}) \\
&\leq \frac{C(\alpha, \beta)}{f(y)^{1-\alpha}} f(x)^{\frac{\beta-k}{\beta}} \leq \frac{C(\alpha, \beta)}{f(y)^{1-\alpha}} |x-y|^{\beta-k},
\end{aligned}$$

as required. This completes the proof for the first statement. Note that (3.2) follows directly from (3.8), since this last expression also holds for all $0 \leq k < \beta$.

Suppose now $f \geq \varepsilon > 0$. It follows immediately from the results that have just been established that $\|f^\alpha\|_{C^\beta} \leq C \|f\|_{\mathcal{H}^\beta} / \varepsilon^{1-\alpha}$. By (3.2),

$$|(f(x)^\alpha)^{(j)}| \leq C \|f\|_{\mathcal{H}^\beta}^{j/\beta} f(x)^{\alpha-j/\beta} (f(x)/\varepsilon)^{\frac{j(1-\alpha)}{\beta}} = C (\|f\|_{\mathcal{H}^\beta} / \varepsilon^{1-\alpha})^{\frac{j}{\beta}} (f(x)^\alpha)^{\frac{\beta-j}{\beta}}$$

for all $1 \leq j < \beta$, so that also $|f^\alpha|_{\mathcal{H}^\beta} \leq C \|f\|_{\mathcal{H}^\beta} / \varepsilon^{1-\alpha}$. \square

Using Lemmas 1, 2 and 3 we can now prove Proposition 1 and Theorem 6.

Proof of Proposition 1. By rescaling, we may assume that $\|f\|_{\mathcal{H}^\beta} = 1$ for both statements. For $\eta > 0$, note that $|f + \eta|_{C^\beta} = |f|_{C^\beta}$ and $|f + \eta|_{\mathcal{H}^\beta} \leq |f|_{\mathcal{H}^\beta}$, since adding a positive constant function to $f \geq 0$ can only reduce the $|\cdot|_{\mathcal{H}^\beta}$ -seminorm. Applying Lemma 3 thus yields $|(f + \eta)^\alpha|_{C^\beta} \leq C/\eta^{1-\alpha}$. Using that $|(x + \eta)^\alpha - x^\alpha| \leq \eta^\alpha$ for all $x \geq 0$ and the wavelet bound for C^β -functions in Lemma 1, we have

$$|\langle f^\alpha, \psi_{j,k} \rangle| \leq \left| \int (f + \eta)^\alpha \psi_{j,k} \right| + \left| \int ((f + \eta)^\alpha - f^\alpha) \psi_{j,k} \right| \leq \frac{C}{\eta^{1-\alpha}} 2^{-\frac{j}{2}(2\beta+1)} + C' \eta^\alpha 2^{-j/2}.$$

Optimizing over η yields $\eta = 2^{-j\beta}$ and thence the first result.

We show that for all $j \geq j(x_0)$, the support of any $\psi_{j,k}$ with $\psi_{j,k}(x_0) \neq 0$ is contained in the set $\{x : f(x_0)/2 \leq f(x)\}$. To see this observe that for any such $\psi_{j,k}$ it holds that $|\text{supp}(\psi_{j,k})| \leq 2^{-j} |\text{supp}(\psi)| \leq a f(x_0)^{1/\beta}$ and so applying Lemma 2 yields that $f(t) \geq f(x_0) - |f(t) - f(x_0)| \geq f(x_0)/2$ for any $t \in \text{supp}(\psi_{j,k})$. Using this and applying Lemma 3, we have that $\|f^\alpha\|_{\mathcal{H}^\beta(\text{supp}(\psi_{j,k}))} \leq C(\beta)/(f(x_0)/2)^{1-\alpha}$, where the first norm refers to f^α restricted to $\text{supp}(\psi_{j,k})$ with the obvious modification of $\|\cdot\|_{\mathcal{H}^\beta}$ to this set. Applying Lemma 1 to such (j, k) then yields the result. \square

Proof of Theorem 6. By the wavelet characterization of the Besov space $B_{\infty\infty}^\beta$ (e.g. p. 362 of [8]) and Proposition 1,

$$\|f^\alpha\|_{B_{\infty\infty}^{\alpha\beta}} \asymp \max_k |\langle f^\alpha, \phi_k \rangle| + \sup_{j \geq 0} 2^{j(\alpha\beta+1/2)} \max_k |\langle f^\alpha, \psi_{j,k} \rangle| \leq C(\phi) \|f\|_\infty^\alpha + C(\psi, \alpha, \beta) \|f\|_{\mathcal{H}^\beta}^\alpha.$$

Note that the right-hand side of (3.2) in Lemma 3 can be written as $C(\|f\|_{\mathcal{H}^\beta}^\alpha)^{\frac{k}{\alpha\beta}} (f(x)^\alpha)^{\frac{\alpha\beta-k}{\alpha\beta}}$. Since this holds for all $1 \leq k \leq \lfloor \alpha\beta \rfloor$, this establishes the second statement. The last statement follows from the first two and the equivalence of the norms $\|\cdot\|_{C^s}$ and $\|\cdot\|_{B_{\infty\infty}^s}$ for non-integer s (Theorem 4.3.2 of [8]). \square

Acknowledgements

We would like to thank Armin Rainer for useful suggestions that led to an improved version of the paper.

References

- [1] BONY, J.-M., BROGLIA, F., COLOMBINI, F., AND PERNAZZA, L. Nonnegative functions as squares or sums of squares. *Journal of Functional Analysis* 232, 1 (2006), 137 – 147.
- [2] BONY, J.-M., COLOMBINI, F., AND PERNAZZA, L. On the differentiability class of the admissible square roots of regular nonnegative functions. In *Phase space analysis of partial differential equations*, vol. 69 of *Progr. Nonlinear Differential Equations Appl.* Birkhäuser Boston, Boston, MA, 2006, pp. 45–53.
- [3] BONY, J.-M., COLOMBINI, F., AND PERNAZZA, L. On square roots of class C^m of nonnegative functions of one variable. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* 9, 3 (2010), 635–644.
- [4] COHEN, A., DAUBECHIES, I., AND VIAL, P. Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.* 1, 1 (1993), 54–81.
- [5] COLOMBINI, F., AND LERNER, N. Une procedure de Calderón-Zygmund pour le problème de la racine k -ième. *Ann. Mat. Pura Appl. (4)* 182, 2 (2003), 231–246.
- [6] DIEUDONNÉ, J. Sur un théorème de Glaeser. *J. Analyse Math.* 23 (1970), 85–88.
- [7] GHISI, M., AND GOBBINO, M. Higher order Glaeser inequalities and optimal regularity of roots of real functions. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* 12, 4 (2013), 1001–1021.
- [8] GINÉ, E., AND NICKL, R. *Mathematical Foundations of Infinite-dimensional Statistical Models*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2016.
- [9] GLAESER, G. Racine carrée d’une fonction différentiable. *Annales de l’institut Fourier* 13, 2 (1963), 203–210.
- [10] HÄRDLE, W., KERKYACHARIAN, G., PICARD, D., AND TSYBAKOV, A. *Wavelets, approximation, and statistical applications*. Springer, New York, 1998.
- [11] JOHNSON, W. P. The curious history of Faà di Bruno’s formula. *Amer. Math. Monthly* 109, 3 (2002), 217–234.
- [12] LENGYEL, P. Racines de fonctions différentiables. *Ann. Inst. Fourier (Grenoble)* 25, 2 (1975), xii, 171–183.

- [13] MACCHIA, R. On roots of differentiable functions. *J. Math. Anal. Appl.* 63, 1 (1978), 112–140.
- [14] MANDAI, T. Smoothness of roots of hyperbolic polynomials with respect to one-dimensional parameter. *Bull. Fac. Gen. Ed. Gifu Univ.*, 21 (1985), 115–118.
- [15] NUSSBAUM, M. Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* 24, 6 (12 1996), 2399–2430.
- [16] PARUSIŃSKI, A., AND RAINER, A. A new proof of Bronshtein’s theorem. *J. Hyperbolic Differ. Equ.* 12, 4 (2015), 671–688.
- [17] PARUSIŃSKI, A., AND RAINER, A. Optimal Sobolev regularity of roots of polynomials. *ArXiv e-prints* (June 2015).
- [18] PARUSIŃSKI, A., AND RAINER, A. Regularity of roots of polynomials . *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* (2016), To appear.
- [19] PATSCHKOWSKI, T., AND ROHDE, A. Adaptation to lowest density regions with application to support recovery. *Ann. Statist.* 44, 1 (2016), 255–287.
- [20] RAINER, A. Smooth roots of hyperbolic polynomials with definable coefficients. *Israel J. Math.* 184 (2011), 157–182.
- [21] RAY, K., AND SCHMIDT-HIEBER, J. Minimax theory for a class of nonlinear statistical inverse problems. *Inverse Problems* 32, 6 (2016), 065003.
- [22] REICHARD, K. Roots of differentiable functions of one real variable. *J. Math. Anal. Appl.* 74, 2 (1980), 441–445.